Adenopathy occurs more commonly in older children and may reflect greater immunological maturity.[6] In one large series 11 of 450 patients with Kawasaki disease presented with cervical lymphadenitis as the primary complaint.[6] The neck nodes were large and often erythematous and tender, although pronounced cellulitis, as seen in our patient, was not described. The diagnosis of Kawasaki disease was delayed in most of the children; three developed coronary artery lesions. None required intubation. The cervical lymphadenopathy and fever resolved promptly in those given intravenous immunoglobulin. Kawasaki disease may also present with upper airway inflammation[7] or mimic a retropharyngeal abscess.[8]

In the United States Kawasaki disease is the leading cause of acquired heart disease in childhood.[3] In most cases cardiac lesions can be prevented by prompt diagnosis and treatment. Not all of the diagnostic features of Kawasaki disease may be obvious on presentation, and one particular sign, such as lymphadenopathy, may predominate. The diagnosis of Kawasaki disease, however, should always be considered in a young child with unexplained and prolonged fever, and careful and repeated examination is needed as the classical features of Kawasaki disease may only become evident over time.

Funding: None.
Conflict of interest: None.

1 Schiller B, Fasth A, Bjorkhem G, Elinder G. Kawasaki disease in Sweden: incidence and clinical features. *Acta Paediatr* 1995;84:769-74.
2 Yanagawa H, Yashiro M, Nakamura Y, Kawasaki T, Kato H. Epidemiological pictures of Kawasaki disease in Japan: from the nationwide incidence survey in 1991 and 1992. *Pediatrics* 1995;95:475-9.
3 Rowley AH, Gonzalez-Crussi F, Shulman ST. Kawasaki syndrome. *Rev Infect Dis* 1988;10:1-15.
4 Hsu CH, Chen MR, Hwang FY, Kao HA, Hung HY, Hsu CH. Efficacy of plasmin-treated intravenous gamma-globulin for therapy of Kawasaki syndrome. *Pediatr Infect Dis J* 1993;12:509-12.
5 Durongpisitkul K, Gururaj VJ, Park JM, Martin CF. The prevention of coronary artery aneurysm in Kawasaki Disease: a meta-analysis on the efficacy of aspirin and immunoglobulin treatment. *Pediatrics* 1995; 96: 1057-61.
6 Stamos JK, Corydon K, Donaldson J, Shulman ST. Lymphadenitis as the dominant manifestation of Kawasaki disease. *Pediatrics* 1994;93:525-8.
7 Kazi A, Gauthier M, Lebel MH, Farrell CA, Lacroix J. Uvulitis and supraglottitis: early manifestations of Kawasaki disease. *J Pediatr* 1992; 120:564-7.
8 Pontell J, Rosenfeld RM, Kohn B. Kawasaki disease mimicking retropharyngeal abscess. *Otolaryngol Head Neck Surg* 1994;110:428-30.

(Accepted 18 March 1996)

Statistics Notes

Comparing several groups using analysis of variance

Douglas G Altman, J Martin Bland

This is the 20th in a series of occasional notes on medical statistics

Many studies, including most controlled clinical trials, contrast data from two different groups of subjects. Observations which are measurements are often analysed by the t test, a method which assumes that the data in the different groups come from populations where the observations have a normal distribution and the same variances (or standard deviations). While the t test is well known, many researchers seem unaware of the correct method for comparing three or more groups. For example, table 1 shows measurements of galactose binding for three groups of patients. A common error is to compare each pair of groups using separate two sample t tests[1] with the consequent problem of multiple testing.[2] The correct approach is to use one way analysis of variance (also called ANOVA), which is based on the same assumptions as the t test. We compare the groups to evaluate whether there is evidence that the means of the populations differ. Why then is the method called analysis of variance?

We can partition the variability of the individual data values into components corresponding to within and between group variation. Table 2 shows the analysis of variance table for the data in table 1. Fuller details about the calculations can be found in textbooks[3] (although a computer would generally be used). The first column shows the "sum of squares" associated with each source of variation; these add to give the total sum of squares. The second column shows the corresponding degrees of freedom. For the comparison of k groups there are k–1 degrees of freedom. The third column gives the sums of squares divided by the degrees of freedom, which are the variances associated with each component (perhaps confusingly called mean squares). When there are two groups the residual variance is the same as the pooled variance used in the two sample t test.

Analysis of variance assesses whether the variability of the group means—that is, the between group variance—is greater than would be expected by chance. Under the null hypothesis that all the population means are the same the between and within group variances will be the same, and so their expected ratio would be 1. The test statistic is thus the ratio of the between and within group variances, denoted F in table 2. The larger the value of F the more evidence there is that the means of the groups differ. The observed value of F is compared with a table of values of the F distribution using the degrees of freedom for both the numerator and denominator—this value is sometimes written as $F^{[2,39]}$. For the data in table 1 an F value greater than 3.24 would be significant with $P<0.05$. The observed value is far larger than this, giving strong evidence that the three populations of patients differ. With two groups one way analysis of variance is exactly equivalent to the usual two sample t test, and we have $F=t^2$.

When the groups are significantly different we will often wish to explore further to see where the differences lie. When we compare more than two groups

Table 1—*Measurements of galactose binding in three groups of patients (data from M Weldon)*

Crohn's disease	Ulcerative colitis	Controls	
1343	1264	1809	2850
1393	1314	1926	2964
1420	1399	2283	2973
1641	1605	2384	3171
1897	2385	2447	3257
2160	2511	2479	3271
2169	2514	2495	3288
2279	2767	2525	3358
2890	2827	2541	3643
	2895	2769	3657
	3011		
	3013		
	3355		
Mean 1910.2	2373.8	2804.5	
SD 515.7	727.1	526.8	

IRCF Medical Statistics Group, Centre for Statistics in Medicine, Institute of Health Sciences, PO Box 777, Oxford OX3 7LF
Douglas G Altman, *head*

Department of Public Health Sciences, St George's Hospital Medical School, London SW17 0RE
J Martin Bland, *professor of medical statistics*

Correspondence to:
Mr Altman.

BMJ 1996;312:1472–3

Table 2—*Analysis of variance table for the data in table 1*

Source of variation	Degrees of freedom	Sum of squares	Mean square	Variance ratio (*F*)	Probability (P)
Between groups	2	5 174 310.0	2 587 155.0	7.34	0.002
Residual (within groups)	39	13 743 776.2	352 404.5		
Total	41	18 918 086.2			

we need a clear idea of which comparisons we are interested in. Very often we are not equally interested in all possible comparisons. Many statistical procedures are available, their appropriateness depending on the question one wishes to answer. One simple method is to use the residual variance as the basis for modified t tests comparing each pair of groups. Here we get: group 1 v group 2, P=0.12; 1 v 3, P=0.0002; 2 v 3, P=0.06. The main difference is thus between groups 1 and 3, as can be seen from table 1. This procedure is an improvement on simply performing three two sample t tests in the first place because we proceed to comparing pairs of groups only if there is evidence of significant variability among all the groups, and also because we use a more reliable estimate of the variance within groups. Investigation of all pairs of groups often does not yield a simple interpretation, which is the price we can pay for not having a specific hypothesis. When the overall F test is not significant it is generally unwise to explore differences between pairs of groups. If the groups have a natural ordering—for example, representing patients with different stages of a disease—it is preferable to examine directly evidence for a (linear) trend in means across the groups.[1] We will consider such data in a subsequent statistics note.

This type of analysis can be extended to more complex data sets with two classifying variables, using two way analysis of variance, and so on. Analysis of variance is a special type of regression analysis, and most data sets for which analysis of variance is appropriate can be analysed by regression with the same results.

1 Godfrey K. Comparing the means of several groups. In: Bailar JC , Mosteller F , eds. *Medical uses of statistics*. 2nd ed. Boston, MA: NEJM Books, 1992: 233-57.
2 Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ* 1995;**310**:170.
3 Armitage P, Berry G. *Statistical methods for research workers*. 3rd ed. Oxford: Blackwell, 1994.

Correction

ABC of Urology: Urological trauma and bladder reconstruction

An editorial error occurred in this article by Chris Dawson and Hugh Whitfield (25 May pp 1352-4). The angiograms on p 1352 were printed upside down.

A MEMORABLE PATIENT

Insomnia as a complication of pleural aspiration

While working in tuberculosis clinics in a remote area of Nepal we saw two patients who had difficulty sleeping after pleural aspiration and a striking case of sleep disturbance in the spouse of a third patient.

We first encountered the problem with a 52 year old man who was found to be sputum positive for acid and alcohol fast bacilli. Despite his weight loss and cough, dyspnoea was his presenting symptom. On examination he had a pleural effusion. After a month of short course chemotherapy his sputum was negative for acid and alcohol fast bacilli, but his dyspnoea had worsened and his pleural effusion was larger.

A litre of yellow fluid was aspirated and this was repeated after two days. His dyspnoea improved, but he complained bitterly that he could not sleep properly, as every time he turned over he was wakened by the sensation of water splashing inside his chest. He complied with treatment, gained weight, and was discharged after six months, his pleural effusion gone. He reported that not only had the sensation ceased, but now he had finished his treatment he was once more able to drink the local spirit, raksi, which helped him sleep.

Some months later we met a 28 year old man with a long history of cervical lymphadenopathy and weight loss. Although persistently sputum negative for acid and alcohol fast bacilli, two members of his household were sputum positive. Twice previously he had defaulted from chemotherapy. He had severe dyspnoea and a large pleural effusion; 1.6 litres of pus were aspirated. Two weeks later a further litre was aspirated. His dyspnoea improved considerably, but he too found that when he changed his sleeping position he was roused by splashing inside his body. His sleep was being interrupted several times each night. He deteriorated, but despite starting second line treatment, he died shortly afterwards.

Although we had noted this problem, our next encounter with it surprised us. A 42 year old man presented with similar signs and symptoms to our first patient. His wife had insisted that he seek help. She was aware of the importance of his symptoms, having been successfully treated for tuberculosis herself. She would prove to be a valuable ally in ensuring his compliance to treatment. After one month his dyspnoea had not improved and his pleural effusion was larger; 1.4 litres of turbid yellow fluid were aspirated.

A week later he felt better. He had not noticed any new problems since aspiration but the consultation was interrupted by his wife's shrill complaint that when he moved in his sleep she was disturbed by a splashing sound emanating from his chest. This had never happened during her own treatment; she had had enough and had a good mind to sleep elsewhere.

It seems probable that the sensation and noise causing insomnia for these unfortunate people were due to the creation of a small pneumothorax during aspiration allowing "splashing" to occur at the air and fluid interface.

Default from treatment is arguably the biggest problem in treating tuberculosis in Nepal. Side effects of treatment must be taken seriously as they may increase default rates, especially if they lead to marital disharmony.—H KLONIN *is a senior registrar in Stoke-on-Trent* and C VICKERY *is a general practitioner in the Shetlands.*

We welcome filler articles of up to 600 words on topics such as *A memorable patient, A paper that changed my practice, My most unfortunate mistake,* or any other piece conveying instruction, pathos, or humour. If possible the article should be supplied on a disk.